Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2014/04/jls.htm
2014 Vol. 4 (S4), pp. 399-407/Nejad and Romouzi
Research Article

INVESTIGATION OF VIDEO STREAMING APPLICATION: A PRACTICAL GUIDANCE

***Arefe Esalat Nejad[1] and Morteza Romouzi[2]**
[1]*Young Researchers and Elite Club, Baft Branch, Islamic Azad University, Baft, Iran*
[2]*Department of Electronic and Computer, Kashan Branch, Islamic Azad University, Kashan, Iran*
**Author for Correspondence*

ABSTRACT
Video streaming applications in wireless camera networks composed by low-end devices are attractive for enabling new pervasive high value services. When complying with the 6LoWPAN standard, the reduced amount of available bandwidth imposes the transmission of low resolution images. We propose a source rate control scheme for streaming video sequences over wireless channels by resorting on a reduced-reference (RR) quality estimation approach. It works as follows: the server extracts important features of the original video, which are coded and sent through the channel along with the video sequence and then exploited at the decoder to compute the actual quality; the observed quality is analyzed to obtain information on the impact of the source rate at the given system configuration; at the receiver, decisions are taken on the optimal source rate to be applied next at the encoder to maximize the quality as perceived at the user-side. The rate is adjusted on a per-window basis to compensate low throughput periods with high-throughput periods so as to avoid abrupt video quality changes, which can be caused by sudden variations in the channel throughput. The use of the RR quality estimation represents the main novelty of the proposed work. Urban traffic control systems have based their technological infrastructure both on advanced analog close-circuit television systems (CCTV) and point-to-point links, providing difficult-to-scale and very expensive systems. The main goal of an urban traffic monitoring system is to capture, send, play and distribute video information from the streets of a certain city to a management center where it is processed by different services.

Keywords: *Video Streaming, Networks, Wireless Channels*

INTRODUCTION
On-line surveillance for safety and security is a major requirement of public transport and other public places to address the modern demands of mobility in major urban areas and to effect improvements in quality of life; environmental protection or traffic control (Foresti, 1998). The surveillance task is a complex one involving computer and communications technologies, management procedures and people. Visual surveillance based on closed-circuit television (CCTV) is an important part of such systems, and is being used in many cities around the world (Gettman *et al.*, 2002; Santini, 2000; Sherwood, 1998). The information collected by these systems is sent to a central Traffic Management Centre (TMC). The TMC refers to the point where a variety of information sources(visual, auditory, traffic, timetabling, etc.) is gathered and used by skilled human operators to control and maintain acceptable levels of security and safety (Trivedi *et al.,* 2000). At the TMC the information is processed in different ways by several applications and services: played back on screen monitors, saved in video tapes and sometimes processed and automatically analyzed for different uses like fast incident detection (without human monitoring), estimation of travel times between several points; lane usage; heavy vehicle count or detailed traffic condition information for public use. In some cases the digital treatment of this video signals is required for the processing of the information and for using artificial vision techniques (Kogut and Trivedi, 2001; Kumar *et al.,* 2005; Menendez *et al.,* 1999; Park *et al.,* 1999). Traffic Management Systems (TMS) based in the utilization of video cameras need high quality video signals. Usually, CCTV techniques have been used for the deployment of such systems, due to the provision of high-quality analog video signals (Foresti *et al.,* 2000; Palau *et al.,* 2005). The basic architecture of a TMS includes a central station TMC, interconnected by some sort of wired or wireless communication network to one or several remote

stations located on the sites under surveillance. Each remote station transmits video and other data (such alarm indications or sensor information) to the central station according to some surveillance policy, and locally controls a set of peripherals (cameras, sensors, actuators, traffic lights, etc.) as required to provide an effective surveillance service. As with many other distributed systems, video surveillance systems (like TMS) are traditionally implemented using proprietary middleware, and hardware (Brett and Trivedi, 2002).

The media streaming system that we have designed and developed has as main concern and distinctive features these two previous objectives, which can be stressed in the following points:

- TCP/IP network with redundancy in the topology configured using Spanning Tree. Enough capacity to hold more than a thousand high-quality video streams from the video cameras.
- MPEG4 compression standard utilization for video streams (Aguigui *et al.,* 2000), allowing the feeding of any kind of video processing service or artificial intelligence applications.
- Multicast features in order to playback the same video stream in more than one terminal, optimizing the consumed bandwidth as the number of cameras or processing services increase.
- Fiber optic network deployed in one of the largest Spanish cities.

Notwithstanding the increase in the transmission rate novel wireless network standards for both wide and local area scenarios, the resulting channels are still highly dynamic, with a bit error rate (BER) that punctuates by orders magnitude in less than a second and with an available bandwidth with continuous sudden variations. This is mainly because of the contention-based nature of common access techniques and because of the abrupt changes in radio interference and packet collisions. It follows that the transmission of real-time data at medium/high bitrates, as it is the case of video streaming applications, is still a challenging task, which requires a continuous monitoring of the channel conditions and the content to be transmitted and to adapt the system parameters accordingly, so as to avoid abrupt communication quality drops.

There are two major advancements with respect to the state of the art:

- We go further in taking into account the quality perceived by the content consumer by assessing the video quality ter channel errors, playback buffer starvation occurrences and error concealment. This is achieved using a reduced-reference (RR) approach that relies on features of the original video that are extracted by the server, sent to the decoder, and used to evaluate the quality of the video as watched by the end-user. Such a consumer-side evaluation is used to build a quality curve varying the coding rate, which is then used to drive future decisions on the streaming. The use of a RR approach is new in this context, since in the past proposals video models are used to predict the relationships of the quality with the coding rate, channel errors and starvation occurrences. This is clarified in the section on related works.
- To compute the playback buffer starvation probability we resort on a heuristic but effective technique for taking into account the channel behavior without the need for major modeling assumptions on the wireless channel and complex procedures for the estimation of model parameters. Accordingly, the time needed to deliver a data frame to the client is monitored during the streaming and used to build the relevant pdf (probability density function). Three alternative aging procedures are proposed to reduce the weight of old samples with respect to the recent ones.

Video Compression and Transmission

State-of-the-art video-surveillance applications are based on a streaming service between a static camera and a remote control point; the service is required to fulfill the temporal constraints usually defined in terms of frame rate. The minimum frame rate required by such applications is about 1 fps. The diagram shown in Figure 1 describes the main functional blocks, and their relations, in standard video-surveillance application based on a streaming service. In this situation, on the transmitter side, the images acquired by a camera network devices are compressed using a video-encoder and fragmented in packets to be transmitted through the wireless network. At the receiver side, the received packets are aggregated in order to decompress the images (using a video- decoder) to be shown on the user screen. As it is discussed in (Petracca *et al.,* 2011), a single stream, at a frame rate of 1 fps and QQ-VGA resolution (i.e., 160_120 pixels), fully saturates the available bandwidth in IEEE802.15.4 networks. In the same conditions, considering the 6LoWPAN overhead, a single stream would not fit the temporal constraints as imposed by the minimum required frame rate.

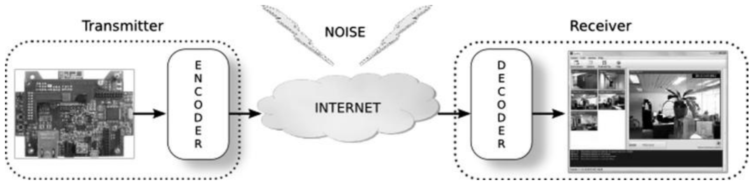

Figure 1: Video-surveillance functional block diagram

System Model

In this section we address the issue of rate control for video streaming applications by considering a typical wireless scenario where the channel resources are shared among a variable number of stations using a contention based access control mechanism. Resource sharing, together with the high variability of traffic load and the fluctuations in bit error rates, brings to bursty channel throughputs as perceived by each end-user contending the channel with the other stations (Tickoo, 2003). To reduce the number of corrupted packets and obtain reliable transmissions, a Forward Error Correction (FEC) scheme is adopted. While FEC coding adds redundancy to the packets so as to correct possible channel errors, errors may still be present at the decoder. In this case, the packet is not discarded but the error detection and resynchronization features of the video codec are adopted (e.g., header and slice synchronization of the MPEG standard). No retransmissions are performed to limit the streaming end-to-end playback delay. Figure 2 depicts the architecture of the considered video streaming system, where Table 1 provides the main notation used throughout the manuscript with the Section where the symbols appear for the rest time.

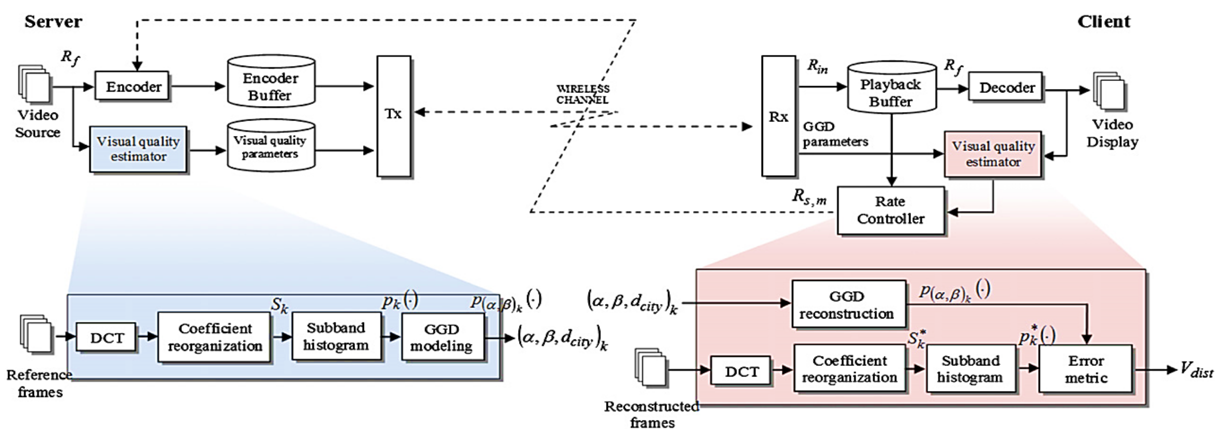

Figure 2: Reference architecture of the video streaming system

The objective is to embed enough information in the video stream so that the quality of the received sequences after decoding (including both the coding and the channel distortions) is computed and used for the estimation of future characteristics. The expected quality, Q_m, as a function of the source coding rate for next window m is expressed as follows:

$$Q_m(R_{s,m}) = [1-\Phi_m(B_m,R_{s,m},H)]Q_m^{nst}(R_{s,m}) + \Phi_m(B_m,R_{s,m},H)Q_m^{st}(R_{s,m}).$$

There are two components in this expression which are related to the probability of starvation at the playback buffer ϕ_m The probability of starvation depends on the status of the playback buffer B_m coding rate $R_{s,m}$ m, on the used source, and on the statistics about the channels, which are provided by the histogram of packet transmission time H. As to the video quality, it depends on the source coding bitrate both in case of uninterrupted decoding, Q_m^{nst}, and in case of starvation, Q_m^{st}.

Table 1: Main notation used in the paper

Notation	Description	Section
R_f	Frame rate	3
R_{in}	Source coding rate	
N	Number of frames accumulated at the beginning (before the playback commences)	
T	Window size in the rate control scheme	
m	Index of the windows in the rate control scheme	
$R_{s,m}$	Source coding rate during window m	
Q_m	Function of the quality expected during window m	
Φ_m	Probability of starvation at the playback buffer at the end of window m	
B_m	Playback buffer occupancy at the beginning of window m	
Q_m^{st}	Function of the quality expected during window m in case of starvation	
Q_m^{nst}	Function of the quality expected during window m in case of no starvation	
S_k	kth DCT subband	4
α,β,d	Reduced-reference quality model parameters to be send to the receiver for each DCT subband of each frame of interest	
V_{dist}	Reduced-reference quality (error) metric	
γ	Weighting factor used to update quality functions Q_m^{st} and Q_m^{nst}	
$B_m(\tau)$	Playback buffer occupancy during window m	5.1
$Y_m(\tau)$	Number of video bit received during window m	
$\Phi_m(\tau)$	Probability of starvation at the playback buffer during window m	
$P_\tau(n)$	Probability to successfully transmit exactly n frames within τ	
δ_i	Time needed to transmit frame i	
$f_\delta(x)$	pdf of the frame transmission time	
$f_{Y_{\tau,m}}(\cdot)$	pdf for process $Y_m(\tau)$	
A	Aging factor	5.2
f_a,t_a	Aging frequency and time	
$H(x)$	Histogram of the observed frame transmission times	
c	Aging coefficient	
Z	Number of wireless On–Off sessions	7.1
$\lambda_{On},\lambda_{Off}$	Parameters of the exponential distributions for the On and Off periods of the source models	
C	Channel capacity	

Network Infrastructure

The most critical feature of the TMS is communications and specifically the capacity of the network, because the amount of information that has to be transmitted through it is devastating. Every TMS component is connected by means of a corporate high-speed LAN. The network has been deployed using optical fiber, and CISCO networking equipment. It has a hierarchical structure, composed by three layers.

The nodes of each layer are connected to more than one node of the superior layer; this structure allows flexibility, scalability and fault tolerance. The three layers are:

• **Core Network:** It is a high-speed TCP/IP QoS and multicast aware Gigabit Ethernet LAN, working over optical fiber links. This technology allows coverage of the whole city, due to the standard links length, with only three switches (see Figure 3a and b). The TMC holds a node which is part of the Core Network, and serves as support for a LAN to distribute the video streams to the different applications and services: screen monitors for video playback; artificial vision server; storage servers; even some still images of the video cameras of some itineraries can be accessed through Internet.

• **Distribution Network:** Is composed by six Fast Ethernet switches, with Gigabit Ethernet ports for the connection with the Core Network (se Figure 3c and d). These switches receive traffic form the Access Network where the video servers are connected. Each distribution node is connected to two Core Network nodes, in order to achieve fault tolerance and overcome link failures.

• **Access Network:** Each access node is connected to a Distribution Network node using a Fast Ethernet link. The switches used to provide access to the private traffic network of the city are located close to the camera group they serve. Video Servers transmit the video streams to the network and are received in the TMC by different applications performing various tasks. Each Access Network switch usually provides service to one video streaming server, which controls a standard camera group composed by five cameras (four fixed and one mobile), although there are several non-standard camera groups with more or less cameras. Traffic sensors and traffic lights managers deployed within the city are also attached to the Access Network.

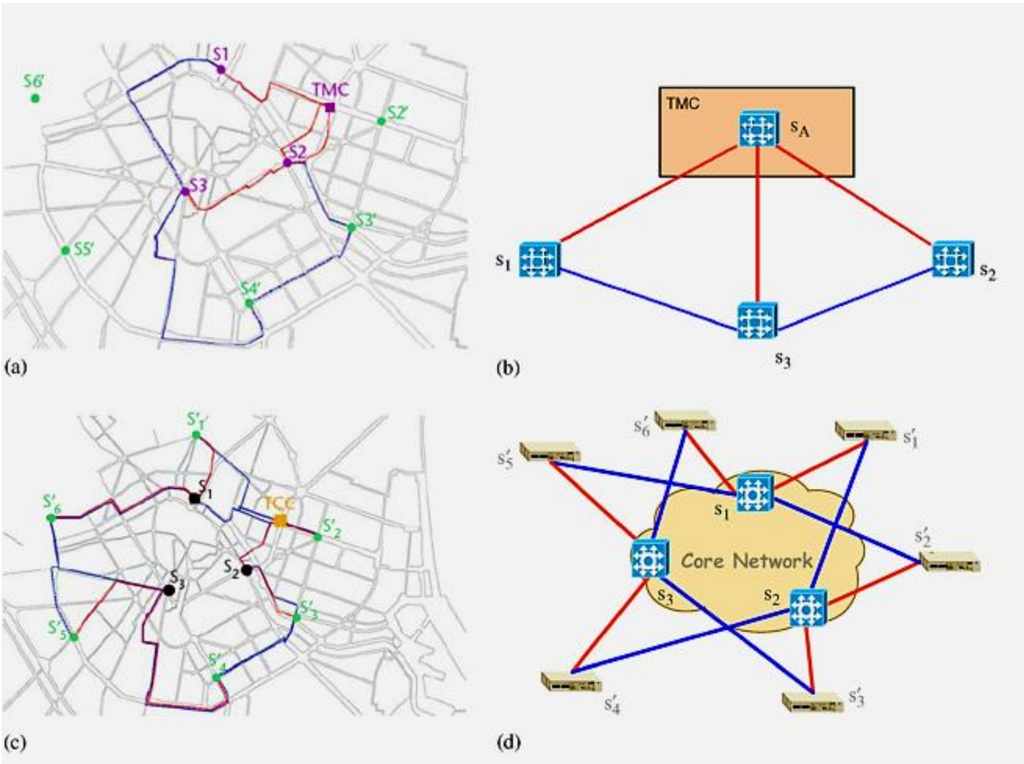

Figure 3: Scheme of the core and distribution network

The video streams nature has some requirements that have influenced the design of the traffic network:

1. **Enough Capacity:** to support as many video streams as possible, at the beginning 600 video streams, finally, when the system will be fully operative, over 1000 flows, with slight modifications of the network.

2. **Multicast Capacity:** due to the TMS structure video streams should have to be directed to more than one playback terminal (e.g. playback monitor and storage servers), and since video streams are high bandwidth consuming applications, even a small number of clients receiving the streams are often enough to saturate the TMS network.
3. **Fault Tolerance:** so more than one route has been considered for each stream to arrive to the TMC. Spanning Tree is executed in order to avoid problems with path duplicity. Fault tolerance is mandatory in any real-time application due to link failure caused by human or non-human factors.

Video Streaming Application

The Video Streaming Application is the upper layer of the TMS. The application is composed of three parts:
1. Video Server side is in charge of capturing the raw video signal, digitalize, compress and transmit each video stream to a multicast destination address. Each camera group is connected to a video server; each of them can manage a maximum of seven streams so with non-standard large groups more than one video server per group would be needed. The destination address is provided in the configuration phase with a very simplistic rule (Palau *et al*., 2005).
2. TMC side, are the group of services executed in different servers in the TMC. Each service listens to several multicast addresses in order to receive and process the video streams sent from the video servers. Several servers in the TMC can receive the same video stream profiting from the multicast paradigm. The services listen to the multicast addresses (i.e. receive the streams) mapped by the Video Streaming Management Application, by means of the management console.
3. Management side distributed within the TMC and the video servers. This part of the Video Streaming Application controls that each screen monitor, storage server and application server receives the adequate video streams, keeping track of the network capacity, server capacity and used multicast addresses. The management console interacts with the Video Streaming Management Server configuring any action, e.g. assigning a playback server to a video server, so the playback server attaches to the corresponding multicast address.

Figure 4 provides an accurate idea of the structure of the video streaming distributed application, and the benefits of using multicast for the distribution of the video streams.

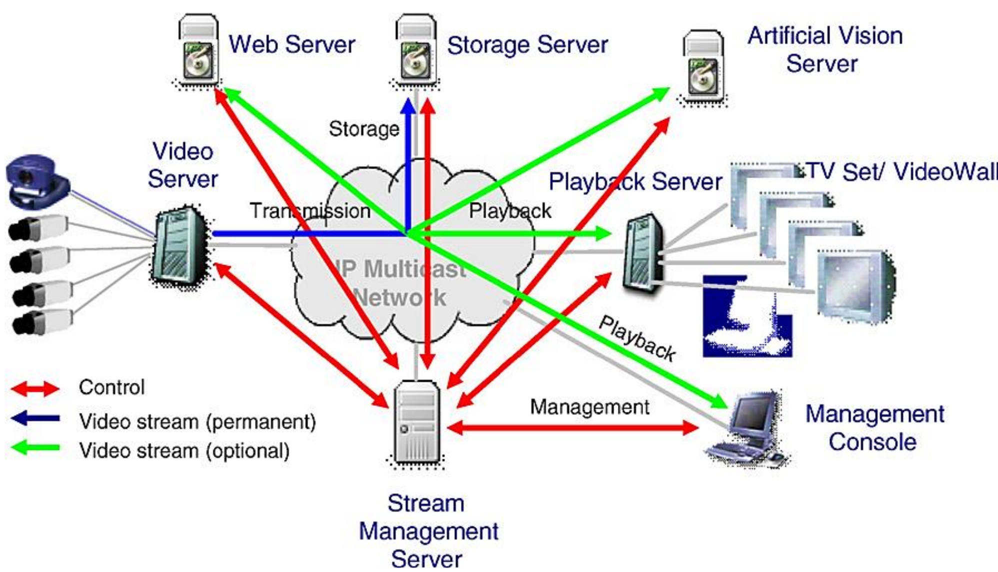

Figure 4: Video streaming application structure

Video Server

The Video Servers are deployed together with camera groups in selected spots in the city streets. Video Servers carry out the digitalization, compression and transmission of the video streams of the cameras connected to them (a video server uses to have five cameras connected and it can deal at most with seven). Each video stream is transmitted to a different multicast address in order to preserve bandwidth and without loss of flexibility as a result of having several services receiving and processing the same stream.

The main tasks carried out by the video servers are the following: (i) raw video capture from analog sources, using standard video acquisition cards; (ii) video compression, usingMPEG4; (iii) packetize the compressed video and attach timestamps within the packets, using MPEG-TS and (iv) transmit packets over the communication network, to the multicast address fixed by the Stream Management Server using UDP.

Multicast Addressing

Each video stream is transmitted to a pre-assigned IP multicast address destination. Due to the TMS requirements, every video stream is continuously saved in the storage server (currently 7 days of continuous high-quality video MPEG4 format), which could be recovered by the application operators if needed, or exported to other applications and systems.

At the same time the video streams are stored other applications and services can process the same video stream, thus each video stream has to be transmitted non-stop, and the multicast addresses cannot be shared. The TMS uses a very simple heuristic to assign the multicast addresses based on two parameters: (i) camera group id, and (ii) camera id within the corresponding group. Each camera group receives a number in order to identify it; currently there are 104 camera groups. The standard camera group is formed by five cameras, although there are smaller and larger groups, e.g. all the cameras installed in a tunnel belong to the same group; it is the largest and comprises 19 cameras.

Overprinting

Video streaming systems used for surveillance or traffic management usually need a labelling mechanism to identify the camera or the video stream in the playback phase. This identification allows direct relationship between the visualized video stream and the camera at the command post (i.e. the TMC) making easier the operation of the users.

Proprietary video servers usually provide this option, and overprint some configurable information over the video images before coding and compressing, securing that the identification data remains with the video during the visualization and the storage processes. The most frequent overprinted data is the id of the camera and the time in the camera position. The id of the camera does not change, but time is a value that changes and evolves. Figure 5 presents the video stream with and without overprinting of the video camera id, 104M01, indicating it is the camera 1 of the 104th group and that it is a mobile camera (M stands for mobile and F for fixed). Under the camera id the date and the hour are overprinted.

Figure 5: Overprinting for playback and storage

Figure 6 represents accurately the process. The metadata are easily recovered from the storage servers because the video streams are saved with the transport headers in the file system.

Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2014/04/jls.htm
2014 Vol. 4 (S4), pp. 399-407/Nejad and Romouzi

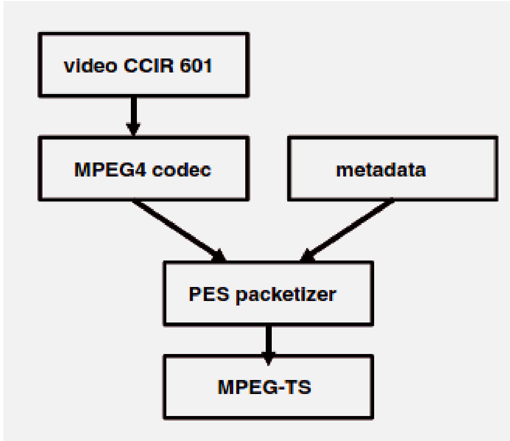

Figure 6: Packetization process

Video Playback

The visualization process is performed in real-time or on-demand accessing the storage server. Real time playback is a simple procedure because the video streams are continuously received at the TMC, the video client only has to join to the corresponding multicast address and start receiving the video stream. On-demand video playback is different because it is performed using unicast mechanisms. When a client wants to playback a saved video clip it is requested using the corresponding application and protocol, accessing to the corresponding storage server.

Video Storage

Multicast video streaming has the advantage of allowing multiple receivers accessing to the same video stream without an increase in the consumed bandwidth. The TMS continuously receives the streams at the TMC. But these streams are not continuously displayed on the monitors, because the most common operation is cyclical playback and even some groups can be assigned neither to any playback device nor to other service or application. Nevertheless video streams are continuously received in the storage servers, in order to save cyclically 7 days of high-quality video that can be recovered for analysis and processing, i.e. video information from cameras in the TMS, can be visualized for real-time but on-demand for the management of past events. Then the main service of the TMS is the storage of the video streams, because this service is always on, any other service like playback is executed in addition to the storage service. The storage equipments are installed at the TMC, currently the system counts with 600high-quality video streams. As in other blocks of the TMS, storage is based in COTS philosophy and PC with massive storage devices have been used and configured to create a Storage Area Network (SAN).

Video Streaming Management Application

The TMS is integrated by video servers, which capture, compress and transmit the video information from the cameras placed on the street, generating the video streams; and the TMC where all the playback, storage and video streaming processing applications reside. Over these two key parts of the TMS, is placed the Video Streaming Management Application (VSMA) which keeps track of the state of all the streams across the network, the receiving applications, the addresses used and controls the storage and publishing of the video data. The VSMA interact with the Management Console executed by the operators.

Remote Control and Mobile Cameras

Remote control information has real time requirements (low delay), because when the operator issues the pan/tilt/zoom operations is better if there is a low delay affecting the playback in order to coordinate the movements. The remote control information is sent to the VSMS and it resends the information to the corresponding video server which is connected to that mobile camera, and is producing the video streams of the group. The system reduces delays to the minimum treating remote control information as high priority data. The video server has been set up to have a low delay in the encoding phase.

REFERENCES

Foresti GL (1998). A real-time system for video surveillance of unattended outdoor environments. IEEE Trans. *Circuits Systems Video Technology* **8**(6) 697–704.

Gettman D (2002). Innovative, ethernet-based communications network architecture for integrated transportation management systems. *In: Proceedings ITS World Congress*, Chicago, IL 1000–8.

Santini S (2000). Analysis of traffic flows in urban areas using web cameras. *In: Proceedings of IEEE Workshop on Applications of Computer Vision, Palm Springs,* CA, USA 140–5.

Sherwood PK (1998). Traffic surveillance and control systems for the area of Hong Kong. *In: Proceedings of International Conference of Road Transport Information and Control* 191–4.

Trivedi MM, Mikic I and Kogut G (2000). Distributed video networks for incident detection and management. *In: Proceedings of IEEE Conference on Intelligent Transportation Systems*, Dearborn, MI, USA 155–60.

Kogut G and Trivedi M (2001). Maintaining the identity of multiple vehicles as they travel through a video network. *In: Proceedings of the IEEE International Conference on Intelligent Transportation Systems*, Oakland, CA, USA 587–93.

Kumar P, Ranganath S, Weimin H and Sengupta K (2005). Framework for real-time behavior interpretation from traffic video. *IEEE Transactions on Intelligent Transportation Systems* **6**(1) 43–53.

Menendez JM, Salgado L, Rendon E and Garcia N (1999). Motorway surveillance through stereo computer vision. *In: Procedings of IEEE Conference on Security Technology* 197–202.

Park SH, Jung K, Hea JK and Kim HJ (1999). Vision-based traffic surveillance system on the internet. *In: Proceedings of the Third International Conference on Computational Intelligence and Multimedia Applications*, New Delhi, India 201–5.

Foresti GL, Mahonen P and Regazzoni CS (2000). *Multimedia Video-based Surveillance Systems: Requirements, Issues and Solutions*. Norwell (Kluwer) MA.

Palau C, Esteve M, Martinez J, Molina B and Perez I (2005). Urban traffic control: a streaming multimedia approach. *In: Proceedings of IEEE ICME*, Amsterdam, Netherlands and July.

Brett T and Trivedi M (2002). A novel interactivity environment for integrated intelligent transportation and telematic systems. *In: Proceedings of the IEEE International Conference on Intelligent Transportation Systems* 340–5.

Aguigui KG, Fehon KJ and Raie R (2000). Overview of the MPEG-4 standard. ISO/IEC JTC1/SC29/WG11 N3444, May/ June.

Petracca M, Ghibaudi M, Salvadori C, Pagano P and Munaretto D (2011). Performance evaluation of FEC techniques based on BCH codes in video streaming over wireless sensor networks, *In: IEEE Symposium on Computers and Communications* 43–48.

Tickoo O and Sikdar B (2003). On the impact of IEEE 802.11 MAC on traffic characteristics, *IEEE Journal on Selected Areas in Communications* **21**(2).

Palau C, Esteve M, Martinez J, Molina B and Perez I (2005). Urban traffic control: a streaming multimedia approach. *In: Proceedings of IEEE ICME*, Amsterdam, Netherlands, and July.